

# MATCHED METRICS TO THE BINARY ASYMMETRIC CHANNELS

CLAUDIO M. QURESHI

**ABSTRACT.** In this paper we establish some criteria to decide when a discrete memoryless channel admits a metric in such a way that the maximum likelihood decoding coincides with the nearest neighbour decoding. In particular we prove a conjecture presented by M. Firer and J. L. Walker establishing that every binary asymmetric channel admits a matched metric.

## 1. INTRODUCTION

As it is well known the Hamming metric is the one that matches to the binary symmetric channel in the sense that maximum likelihood decoding (MLD) coincides with nearest neighbourhood decoding (NND). This type of matching have been studied in other cases, for example in [1] it was characterized certain channels matching to the Lee metric. In that paper, following J. L. Massey [2], a metric is considered matched to a discrete memoryless channel if the NND with respect to this metric is also a MLD. The problem of matching a metric to a channel is taken up by G. Séguin in [3] where the main focus was on sequence of additive metrics. In that paper it is used a stronger condition also assumed here: a metric matched to a channel is one for which not only every NND is a MLD but also every MLD is also a NND. The author obtain necessary and sufficient conditions to the existence of additive metrics matched to a channel and raises the question of what happens if we remove the restriction of the metric be additive. There were not significant progress until the paper [4] by M. Firer and J. Walker where the authors prove, among other results, the existence of a metric (not necessarily additive) matched to the Z-channels and to the  $n$ -fold binary asymmetric channel (BAC) for  $n = 2, 3$  and conjecture that this is also true for  $n > 3$ . Some recent progress in this direction was done in [5] where it is presented an algorithm to decide if a channel is metrizable and in that case return a metric matched to the channel, and in [6] where the author proves that the BAC channels are metrizable in the weaker sense of J. L. Massey in [2].

The main results of this paper are Theorem 4.8, which establishes a necessary and sufficient condition of metrizability of a channel in terms of graph theory, and Theorem 5.2 which proves that the BAC channels are metrizable. Other contributions are the association of channels with graphs (which allows to use the machinery of graph theory to approach problems related to channels) and the introduction of a new structure: the colored posets, which may also be useful in other contexts. This work is organized as follows: In Section 2 we give a brief review of definitions and concepts needed in the development of the paper. In Section 3 we associate a graph with each channel and discuss some results of [4] and [6] in terms of this

graph. In Section 4 it is introduced the concept of colored poset which is used it to describe an algorithm for constructing a metric matched to a channel whenever it is metrizable. A necessary and sufficient criterion for metrizability of a channel is also derived. In Section 5 this criterion is used to prove that the BAC channels are metrizable. In Section 6 we introduce the concept of order of metrizability of a channel and settle some problems related to this.

## 2. PRELIMINARIES

We summarize here some concepts and results to be used in the following sections.

A discrete memoryless channel (simply referred as channel in this paper)  $W : \mathcal{X} \rightarrow \mathcal{X}$  is characterized by its transition matrix related to the input and output alphabet  $\mathcal{X} = \{x_1, x_2, \dots, x_N\}$ . This matrix  $[W] \in \mathcal{M}_N(\mathbb{R})$  is given by  $[W]_{ij} = \Pr_W(x_i|x_j)$ , the probability of receiving  $x_i$  if  $x_j$  was sent. When the channel is understood, this conditional probability is denoted simply by  $\Pr(x_i|x_j)$ .

In this paper we only deal with reasonable channels (in the sense of [4]) and we refer to it simply as channels. A reasonable channel  $W : \mathcal{X} \rightarrow \mathcal{X}$  is one verifying

$$(1) \quad \Pr(x|x) > \Pr(x|y), \quad \forall x, y \in \mathcal{X} \text{ with } y \neq x,$$

which is a necessary condition for the channel to be metrizable.

A channel  $W : \mathcal{X} \rightarrow \mathcal{X}$  is metrizable (in the stronger sense of [4] and [3]) if there is a metric  $d : \mathcal{X} \times \mathcal{X} \rightarrow [0, \infty)$  such that every nearest neighbourhood decoder is a maximum likelihood decoder and vice versa. This is also equivalent to each of the following statements:

- i) For all  $x \in \mathcal{X}$  and every code  $C \subseteq \mathcal{X}$  we have

$$\arg \max_{y \in C} \Pr(x|y) = \arg \min_{y \in C} d(x, y),$$

where both  $\arg \max$  and  $\arg \min$  are interpreted as returning list of size at least 1.

- ii) For all  $x, y, z \in \mathcal{X}$  the following condition holds:

$$\Pr(x|y) \leq \Pr(x|z) \Leftrightarrow d(x, y) \geq d(x, z),$$

for all  $x, y, z \in \mathcal{X}$ .

We remark that the condition (ii) does not imply (i) if the hypothesis of the channel to be reasonable is omitted.

Let  $W : \mathcal{X} \rightarrow \mathcal{X}$  be a channel and  $f : \mathcal{X}^{(2)} \rightarrow [0, +\infty)$  be a function where  $\mathcal{X}^{(2)} = \{A \subseteq \mathcal{X} : \#A = 2\}$  (i.e. the elements of  $\mathcal{X}^{(2)}$  are the 2-subsets of  $\mathcal{X}$ ). We say that  $f$  is coherent with  $W$  if

$$\Pr(x|y) \leq \Pr(x|z) \Leftrightarrow f(x, y) \leq f(x, z),$$

for all  $x, y, z \in \mathcal{X}$  with  $x \neq y$  and  $x \neq z$ . The function is called symmetric when  $f(x, y) = f(y, x)$  for all  $x, y \in \mathcal{X}$  with  $x \neq y$ . If there exists a symmetric function coherent with  $W$  then the channel is metrizable (Lemma 7 of [4]). More explicitly,

if  $f$  is a coherent-with- $W$  symmetric function and we denote by  $N = \max\{f(x) : x \in \mathcal{X}^{(2)}\}$ , then the function  $d : \mathcal{X} \times \mathcal{X} \rightarrow [0, +\infty)$  defined by

$$(2) \quad d(x, y) = \begin{cases} 3N - f(x, y) & \text{if } x \neq y \\ 0 & \text{if } x = y \end{cases}$$

is a metric matched to the channel  $W$ .

The binary (1-fold) asymmetric channel with parameters  $(p, q) \in [0, 1]^2$  (denoted by  $BAC^1(p, q)$ ) is the channel with input and output alphabet  $\mathbb{Z}_2 = \{0, 1\}$  and conditional probabilities  $\Pr_1(1|0) = p$  and  $\Pr_1(0|1) = q$  (and  $\Pr_1(0|0) = 1 - p$  and  $\Pr_1(1|1) = 1 - q$ ). The  $n$ -fold binary asymmetric channel  $BAC^n(p, q)$  is the channel with input and output alphabet  $\mathcal{X} = \mathbb{Z}_2^n$  and for  $x = (x_1, \dots, x_n)$  and  $y = (y_1, \dots, y_n)$  in  $\mathbb{Z}_2^n$  the conditional probabilities are given by

$$\Pr(x|y) = \prod_{i=1}^n \Pr_1(x_i|y_i).$$

We remark that the channel  $BAC^n(p, q)$  verifies condition (1) if and only if  $p + q < 1$  (therefore only this case will be considered in this paper). Indeed, for  $n = 1$  it is obvious and for  $n > 1$  it is a direct consequence of (4) in Section 5. The metrizable of  $BAC^n(p, q)$  is established in [4] for the case  $pq = 0$  and  $n$  arbitrary (the  $n$ -fold  $Z$ -channel) and for  $p + q < 1$  and  $n = 2, 3$ .

A partially ordered set (or simply a poset) is a pair  $(P, \leq)$  where  $\leq$  is a partial order relation (i.e. verifies reflexivity, antisymmetry and transitivity). As usual, when the context is clear the poset is denoted simply by  $P$ . From a directed acyclic graph  $G = (P, E)$  we obtain a natural poset structure on  $P$  by defining  $x \leq y$  if  $x = y$  or there is a (directed) path from  $x$  to  $y$ .

### 3. THE GRAPH $\mathcal{G}_1$ ASSOCIATED WITH A CHANNEL

We associate with each channel (given by its transition matrix) a graph which play an important role in the proof of the metrization of the BAC channel.

**Definition 3.1.** Let  $W : \mathcal{X} \rightarrow \mathcal{X}$  be a channel. The digraph associated with  $W$  (denoted by  $\mathcal{G}_1(W)$ ) has as vertex set  $\mathcal{X}^{(2)}$  (the 2-subsets of  $\mathcal{X}$ ) and directed edges linking  $\{i, j\}$  to  $\{i, k\}$  when  $\Pr(i|j) < \Pr(i|k)$ .

A necessary condition to guarantee the non-existence of a metric matched to a given channel  $W$  is given in Proposition 5 of [4]. This condition states that if the channel  $W : \mathcal{X} \rightarrow \mathcal{X}$  admits a decision chain of length  $r \geq 3$  it is not metrizable. A decision chain of length  $r$  is a sequence  $x_0, x_1, \dots, x_{r-1} \in \mathcal{X}$  verifying  $\Pr(x_i|x_{i-1}) < \Pr(x_i|x_{i+1})$  for  $i : 0 \leq i < r$ , where the indices are taken modulo  $r$  (we note that the definition given in [4] in terms of  $t$ -decision region is equivalent to the one given here). In terms of the graph  $\mathcal{G}_1(W)$ , Proposition 5 of the referred paper can be rewritten as follows.

**Proposition 3.2.** *If a channel  $W$  is metrizable then their associated graph  $\mathcal{G}_1(W)$  is acyclic.*

The following example shows that the converse is false.

**Example 3.3.** Consider the channel  $W : \mathcal{X} \rightarrow \mathcal{X}$  where  $\mathcal{X} = \{0, 1, 2\}$  with transition matrix  $[W] = \begin{pmatrix} 1/2 & 1/4 & 1/4 \\ 1/4 & 5/12 & 1/3 \\ 1/4 & 1/6 & 7/12 \end{pmatrix}$ . The graph  $\mathcal{G}(W)$  has only two edges  $\{0, 1\} \rightarrow \{1, 2\}$  and  $\{1, 2\} \rightarrow \{0, 2\}$  therefore it is acyclic. However a metric compatible with  $W$  should verify  $d(0, 1) = d(0, 2) = d(2, 0) > d(2, 1) = d(1, 2) > d(1, 0) = d(0, 1)$  which is impossible, therefore  $W$  is not metrizable.

**Definition 3.4.** Let  $W$  be a channel. The graph  $\mathcal{G}_1(W)$  is transitive if for every path  $v_0 \rightarrow v_1 \rightarrow \dots \rightarrow v_{r-1}$  with  $\#(v_0 \cap v_{r-1}) = 1$ , the edge  $v_0 \rightarrow v_{r-1}$  belongs to  $\mathcal{G}$ .

It is easy to see that if  $\mathcal{G}_1(W)$  is transitive then it is acyclic. The converse is false (the same channel of Example 3.3 works as counterexample). Proposition 3.2 can be generalized as follows.

**Proposition 3.5.** *If a channel  $W$  is metrizable then its associated graph  $\mathcal{G}_1(W)$  is transitive.*

This proposition has easy verification and can be obtained as a particular case of Theorem 4.8 in Section 5. Since every directed acyclic graph can be associated with a poset (as mentioned in Section 2), then we can associate a poset to a channel whenever its associated graph is acyclic. The transitivity of the graph  $\mathcal{G}_1(W)$  means that if  $v < w$  and  $v \cap w \neq \emptyset$  then  $v \rightarrow w$  is an edge of  $\mathcal{G}_1(W)$ .

Let  $W : \mathcal{X} \rightarrow \mathcal{X}$  be a channel. In terms of the conditional probabilities of  $W$ , the condition for the graph  $\mathcal{G}_1(W)$  to be transitive can be written as follows:  $\mathcal{G}_1(W)$  is transitive if and only if every sequence  $x_0, x_1, \dots, x_{r-1} \in \mathcal{X}$  ( $r \geq 3$ ) satisfying  $x_0 \neq x_{r-1}$  and  $\Pr(x_i | x_{i-1}) < \Pr(x_i | x_{i+1})$  for  $0 \leq i \leq r-2$  (indices taken modulo  $r$ ) and  $x_0 \neq x_{r-2}$  also satisfy  $\Pr(x_{r-1} | x_0) < \Pr(x_{r-1} | x_{r-2})$ . This is exactly the condition proposed in [6] to guarantee the existence of a metric  $d$  such that

$$(3) \quad \arg \max_{y \in C} \Pr_W(x|y) \supseteq \arg \min_{y \in C} d(x, y),$$

for all  $C \subseteq \mathcal{X}$  and  $x \in \mathcal{X}$  (interpreting both  $\arg \max$  and  $\arg \min$  as returning list of size at least 1). Using this condition the author also proves that the BAC channels admit a metric verifying (3). The reciprocal of Proposition 3.5 is also false (in other words it is not possible to prove equality in equation (3) under the hypothesis of transitivity).

**Example 3.6.** Let  $W : \mathcal{X} \rightarrow \mathcal{X}$  be the channel with alphabet  $\mathcal{X} = \{0, 1, 2, 3\}$

and matrix transition  $[W] = \begin{pmatrix} 4/9 & 2/9 & 2/9 & 1/9 \\ 2/9 & 4/9 & 2/9 & 1/9 \\ 3/10 & 1/5 & 2/5 & 1/10 \\ 1/8 & 1/8 & 1/4 & 1/2 \end{pmatrix}$ . The graph  $\mathcal{G}(W)$  is

transitive and its Hasse diagram is showed in Figure 1. Every compatible metric should verify  $d(2, 1) > d(2, 0) = d(0, 2) = d(0, 1) = d(1, 0) = d(1, 2) = d(2, 1)$  which is impossible, then  $W$  is not metrizable.

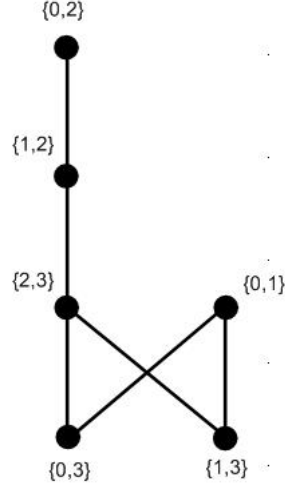


FIGURE 1. The Hasse diagram for the transitive graph  $\mathcal{G}_1(W)$  associated with the non-metrizable channel  $W$ .

#### 4. COLORED POSETS AND A NECESSARY AND SUFFICIENT CONDITION FOR THE CHANNEL TO BE METRIZABLE

Associated with a poset we have a Hasse diagram, which is a representation of the poset in such a way that if  $x < y$  the element  $y$  is above of  $x$  and there is a segment connecting these points whenever there is not  $z$  such that  $x < z < y$ . A height function relates to each element of a poset  $P$ , the maximum possible length of a chain ending in such element. In this paper we refer to this function as the standard height of  $P$  and a height function is any function  $h : P \rightarrow \mathbb{N}$  verifying  $h(x) < h(y)$  whenever  $x < y$ . A subset  $X \subseteq P$  is called horizontal (with respect to  $h$ ) when the restriction of  $h$  to  $X$  is constant. We say that a height is complete when its image is of the form  $[k] = \{0, 1, \dots, k-1\}$  for some  $k \in \mathbb{Z}^+$ . To each height function  $h$  we can associate a Hasse diagram such that  $y$  is above of  $x$  if and only if  $h(y) > h(x)$ . This association establish a bijection between complete heights and Hasse diagrams.

As remarked in the previous section, when the graph  $\mathcal{G}_1(W)$  associated with a channel  $W : \mathcal{X} \rightarrow \mathcal{X}$  is acyclic, we can associate with it a poset  $P = P(W)$  in the set  $\mathcal{X}^{(2)}$ . The standard height verify  $h(\{x, y\}) < h(\{x, z\})$  when  $\Pr(x|y) < \Pr(x|z)$  and then we can easily construct a metric  $d$  with the same property (assuming that the channel verifies (1)). In particular this metric verify (3) and will be weakly metrizable in the sense of [6]. But this metric does not necessarily will match with the channel, since the poset structure of  $W$  (when  $\mathcal{G}_1(W)$  is transitive) does not give information about when two 2-subsets  $\{x, y\}$  and  $\{x, z\}$  verify  $\Pr(x|y) = \Pr(x|z)$  or not, except when they are connected in  $\mathcal{G}_1$ . We need a more general structure to manage also with these cases.

Let  $P$  be a poset. A coloration for  $P$  is any function  $c : P \rightarrow C$  ( $C$  is a finite set) verifying that if  $x < y$  then  $c(x) \neq c(y)$ . A subset  $X \subseteq P$  is monochromatic (with respect to  $c$ ) if the restriction of  $c$  to  $X$  is constant. By definition, every monochromatic set is an antichain of  $P$ . A colored cycle is any sequence  $(x_0, x_1, \dots, x_r)$  in  $P$  verifying that  $x_0 = x_r$  and  $x_i < x_{i+1}$  whenever  $c(x_i) \neq c(x_{i+1})$  for  $0 \leq i < r$ .

**Definition 4.1.** A colored poset is a pair  $(P, c)$  where  $P$  is a poset and  $c$  a coloration for  $P$  such that every colored cycle in  $P$  is monochromatic.

We remark that a colored cycle is not necessarily a cycle in the digraph associated with  $P$ . A Hasse diagram for a colored poset  $(P, c)$  is a Hasse diagram for  $P$  with the additional property that if two points have the same color they are in the same level. The next proposition guarantee the existence of a Hasse diagram for colored posets.

**Proposition 4.2.** *Let  $(P, c)$  be a colored poset. There exists a height function  $h$  for  $P$  such that every monochromatic subset of  $P$  is horizontal with respect to  $h$ .*

*Proof.* We consider the equivalent relation  $x \sim y$  when  $c(x) = c(y)$  and denote by  $P/c = \{A_1, A_2, \dots, A_\kappa\}$  its quotient set. For  $X, Y \subseteq P$  we write  $X \leq Y$  whether  $X = Y$  or there exists  $x \in X$  and  $y \in Y$  such that  $x < y$  (non excluding or). This relation induces a partial order on  $P/c$ . Indeed, the reflexive property is obvious. To check the antisymmetry we assume by the contrary that there exists  $X, Y \in P/c$  such that  $X \leq Y, Y \leq X$  and  $X \neq Y$ . In this case there exists  $x, x' \in X$  and  $y, y' \in Y$  such that  $x \leq y$  and  $y' \leq x'$ . Thus, we would have a non-monochromatic colored cycle  $(x, y, y', x', x)$  ( $c(y) \neq c(y')$  since  $Y$  is an antichain) which is a contradiction, then the relation is antisymmetric. If  $X, Y, Z \in P/c$  satisfies  $X < Y, Y < Z$  and  $X$  and  $Z$  are comparable then  $X < Z$ . Indeed,  $X \neq Z$  because the antisymmetry, if  $X > Z$  we could take  $x, x' \in X, y, y' \in Y$  and  $z, z' \in Z$  such that  $x \leq y, y' \leq z$  and  $z' \leq x'$  and construct the non-monochromatic colored cycle  $(x, y, y', z, z', x', x)$  which is a contradiction. Consider the transitive closure of this relation we obtain a partial order in  $P/c$  and we can sorted the elements of  $P/c$  in such a way that  $A_i < A_j$  imply  $i < j$ . For  $x \in P$  and  $A \in P/c$  we write  $x \geq A$  if there exists some  $a \in A$  such that  $x \geq a$  (otherwise we write  $x \not\geq A$ ).

Now, we construct inductively a sequence  $h_1, h_2, \dots, h_\kappa$  of heights for  $P$  such that  $A_i$  is horizontal with respect to  $h_i$  if  $j \leq i$ . We start considering the standard height  $h_0$  (i.e.  $h_0(x)$  is the length of the largest chain ending in  $x$ ) and define  $h_1$  as follows. Let  $t_1 = \max\{h_0(x) : x \in A_1\}$  and we define:

$$h_1(x) = \begin{cases} \max\{h_0(x) + t_1 - h_0(a) : a \in A_1, a \leq x\}, & \text{if } x \geq A_1, \\ h_0(x), & \text{otherwise.} \end{cases}$$

This function is a height function for  $P$  and  $A_1$  is horizontal with respect to  $h_1$ . Indeed, let us assume that  $x > y$  ( $x, y \in \mathcal{X}^{(2)}$ ) and we consider three cases: (i)  $x > y \geq A_1$ , (ii)  $x \geq A_1$  and  $y \not\geq A_1$  and (iii)  $x, y \not\geq A_1$ . In the first case, since for every  $a \in A_1$  with  $a \leq y$  we have  $a \leq x$  and  $h_0(y) + t_0 - h_0(a) < h_0(x) + t_0 - h_0(a) \leq h_1(x)$ , then  $h_1(y) < h_1(x)$ . In the case (ii), since  $t_1 - h_0(a) \geq 0$  for all  $a \in A_1$  we have  $h_1(y) = h_0(y) < h_0(x) \leq h_1(x)$ . And in the last case  $h_1(y) = h_0(y) < h_0(x) = h_1(x)$ . For all cases, if  $y < x$  then  $h_1(y) < h_1(x)$  and for all  $a \in A_1$ , since  $A_1$  is an antichain we have  $h_1(a) = h_0(a) + t_1 - h_0(a) = t_1$  does not depend on  $a \in A_1$ , so  $A_1$  is horizontal with respect to  $h_1$ . Now we assume that

there exists a height function  $h_m$  for which  $A_1, \dots, A_m$  are horizontal ( $1 \leq m < \kappa$ ) and let  $t_{m+1} = \max\{h_m(x) : x \in A_{m+1}\}$ . We define:

$$h_{m+1}(x) = \max\{h_m(x) + t_{m+1} - h_m(a) : a \in A_m, a \leq x\}$$

if  $x \geq A_m$  and  $h_{m+1}(x) = h_m(x)$  otherwise. Using a similar argument to the case  $m = 1$  (considering three cases) we can prove that  $h_{m+1}$  is a height function (i.e.  $h_{m+1}(x) > h_{m+1}(y)$  whenever  $x > y$ ). Since  $A_{m+1}$  is an antichain, then  $h_{m+1}(a) = t_{m+1}$  for all  $a \in A_{m+1}$ . Let  $x \in A_i$  with  $1 \leq i \leq m$ . We have  $x \not\geq A_{m+1}$  because otherwise we would have  $A_i \geq A_{m+1}$  with  $i < m + 1$  which is a contradiction. Therefore  $h_{m+1}(a) = h_m(a)$  does not depend on  $a \in A_i$  by inductive hypothesis. In the last step (when  $m = \kappa$ ) we obtain a height function  $h_\kappa$  for which all the elements of  $P/c$  are horizontal. In particular, since every monochromatic subset is contained in some  $A_i$ , it is also horizontal with respect to  $h_\kappa$ .  $\square$

**Remark 4.3.** Since the proof of Proposition 4.2 is constructive, it brings us an algorithm to construct a Hasse diagram for a colored poset  $(P, c)$ . We start with the standard height function of  $P$  and after at most  $\kappa$  steps we obtain a height function which induces a Hasse diagram for  $(P, c)$ , where  $\kappa$  is the number of colours used. We can ignore colors which have a unique representative (i.e. unitary equivalence class in  $P/c$ ), since these sets are always horizontal by definition. However, it is important to remark that we cannot ignore horizontal classes with respect to the standard height function except for the singletons and the minimal classes (respect to the order considered in the proof of Proposition 4.2).

**Example 4.4.** Consider the colored poset  $(P, c)$  where  $P = \mathbb{Z}_2 \times \mathbb{Z}_3$  with the order induced by  $00 < 01 < 02, 10 < 11 < 12$  and  $10 < 01$ ; and the coloration  $c : P \rightarrow \{R, B, G, D\}$  given by  $c(00) = c(11) = R$ ,  $c(02) = c(12) = B$ ,  $c(01) = G$  and  $c(10) = D$ . The process for obtaining a Hasse diagram for this colored poset is illustrated in Figure 2. At the final stage we obtain the height function  $h_2 : P \rightarrow \mathbb{N}$  given by  $h_2(10) = 0, h_2(00) = h_2(11) = 1, h_2(01) = 2$  and  $h_2(02) = h_2(12) = 3$ .

Our next goal is to associate with each channel (under certain conditions) a colored poset and to construct a metric from a height function for its Hasse diagram. We start by introducing some graphs associated with a channel.

**Definition 4.5.** Let  $W : \mathcal{X} \rightarrow \mathcal{X}$  be a channel and  $\mathcal{X}^{(2)}$  be the family of 2-subsets of  $\mathcal{X}$ . The graph  $\mathcal{G}(W)$  has as vertices the elements of  $\mathcal{X}^{(2)}$  and the directed edges linking  $\{x, y\}$  to  $\{x, z\}$  if  $y \neq z$  and  $\Pr(x|y) \leq \Pr(x|z)$ . The graph  $\mathcal{G}_0(W)$  is a non-directed graph whose vertices are the elements of  $\mathcal{X}^{(2)}$  and two vertices  $\{x, y\}$  and  $\{x, z\}$  are connected by an edge in  $\mathcal{G}_0(W)$  if  $y \neq z$  and  $\Pr(x|y) = \Pr(x|z)$ .

The graph  $\mathcal{G}_0(W)$  can be identify with a subgraph of  $\mathcal{G}(W)$  with vertices  $\mathcal{X}^{(2)}$  and  $\{x, y\}$  is an edge in  $\mathcal{G}_0(W)$  if and only if  $(x, y)$  and  $(y, x)$  are edges in  $\mathcal{G}(W)$ . By construction the graphs  $\mathcal{G}_0(W)$  and  $\mathcal{G}_1(W)$  have not edges in common, therefore each connected component of  $\mathcal{G}_0(W)$  is an antichain with respect to the order in  $\mathcal{X}^{(2)}$  induced by  $\mathcal{G}_1(W)$  (this fact is used in the next lemma).

**Definition 4.6.** A graph  $G$  is cycle-reverter if for each cycle  $c : v_0 \rightarrow v_1 \rightarrow v_2 \rightarrow \dots \rightarrow v_{r-1} \rightarrow v_0$  in  $G$ , then the reverse cycle  $\bar{c} : v_0 \rightarrow v_{r-1} \rightarrow v_{r-2} \rightarrow \dots \rightarrow v_1 \rightarrow v_0$  is also in  $G$ .

We remark that if  $\mathcal{G}(W)$  is cycle-reverter then the graph  $\mathcal{G}_1(W)$  is transitive (the converse is false) and, in particular, acyclic.

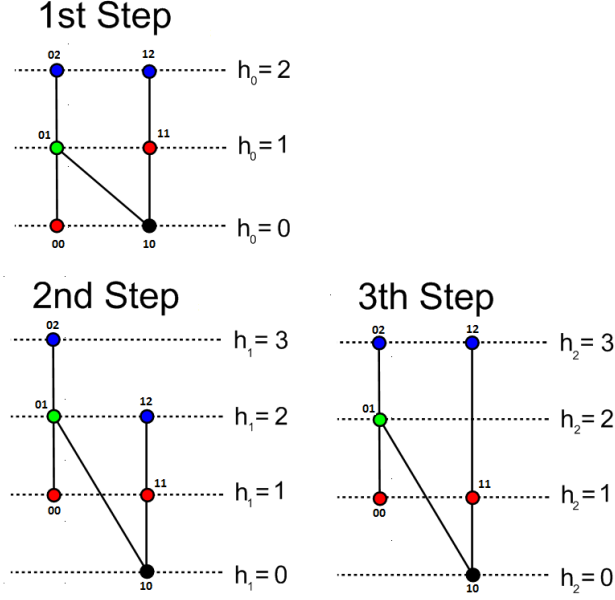


FIGURE 2. The construction of a Hasse diagram for a colored poset.

**Lemma 4.7.** *Let  $W : \mathcal{X} \rightarrow \mathcal{X}$  be a channel such that its associated graph  $\mathcal{G}(W)$  is cycle-reverter and let  $\mathcal{A} = \{A_1, \dots, A_\kappa\}$  be the set of connected components of  $\mathcal{G}_0(W)$ . We consider in  $P = \mathcal{X}^{(2)}$  the poset structure induced by the graph  $\mathcal{G}_1(W)$  (i.e. if  $\{x, y\} \rightarrow \{x, z\}$  is an arrow in  $\mathcal{G}_1$  then  $\{x, y\} < \{x, z\}$ ) and  $c : P \rightarrow \mathcal{A}$  given by  $c(v) = A$  if  $v \in A$ . Then  $(P, c)$  is a colored poset.*

*Proof.* First we prove that  $c$  is a coloration for  $P$ . Let  $v, w \in P$  such that  $v < w$ , then there exists a path  $p_1 : v_0 = v \rightarrow v_1 \rightarrow \dots \rightarrow v_r = w$  (with  $r \geq 1$ ) in  $\mathcal{G}_1(W)$ . In particular  $v = \{x, y\}$  and  $v_1 = \{x, z\}$  with  $\Pr(x|y) < \Pr(x|z)$ . We suppose, to the contrary, that  $c(v) = c(w)$ . Since  $v$  and  $w$  belong to the same connected component in  $\mathcal{G}_0(W)$  there is a path  $p_2$  from  $w$  to  $v$  in  $\mathcal{G}_0(W)$ . Since  $\mathcal{G}(W)$  is cycle-reverter, the reverse of the cycle  $p_1 * p_2$  is also in  $\mathcal{G}(W)$ . In particular the arrow  $v_1 \rightarrow v$  is in  $\mathcal{G}(W)$ , then  $\Pr(x|z) \geq \Pr(y|z)$  which is a contradiction. Now we prove that  $(P, c)$  is a colored poset. We assume, consider a colored cycle  $C = (v_0, v_1, \dots, v_r)$  with  $v_0 = v_r$ . For  $i : 0 \leq i < r$  there is a path  $p_i$  from  $v_i$  to  $v_{i+1}$  in  $\mathcal{G}(W)$  (if  $v_i < v_{i+1}$  this path is in  $\mathcal{G}_1(W)$  and if  $c(v_i) = c(v_{i+1})$  this path is in  $\mathcal{G}_0(W)$ ). Since the graph  $\mathcal{G}(W)$  is cycle-reverter the reverse of the cycle  $p_1 * p_2 * \dots * p_{r-1}$  is a cycle in  $\mathcal{G}(W)$  therefore none of the paths  $p_i$  can be in  $\mathcal{G}_1(W)$  and we conclude that all the vertices  $v_i$  belong to the same connected component in  $\mathcal{G}_0(W)$ , then  $C$  is monochromatic.  $\square$

The next theorem establishes a necessary and sufficient condition for the existence of a metric matched to a given channel.

**Theorem 4.8.** *Let  $W : \mathcal{X} \rightarrow \mathcal{X}$  be a channel. The graph  $\mathcal{G}(W)$  is cycle-reverter if and only if the channel  $W$  is metrizable.*



*Proof.* First we suppose the existence of a metric  $d : \mathcal{X} \times \mathcal{X}$  matched to  $W$  and consider a cycle  $c : v_0 \rightarrow v_1 \rightarrow v_2 \rightarrow \dots \rightarrow v_{r-1} \rightarrow v_0$  in  $\mathcal{G}(W)$ . Every vertex is of the form  $v_i = \{x_i, y_i\}$  with  $x_i, y_i \in \mathcal{X}$  and  $\#(v_i \cap v_{i+1}) = 1$  for  $0 \leq i < r$  (indices taken modulo  $r$ ). Since  $d$  matches with  $W$  from the cycle  $c$  we obtain the following chain of inequalities:

$$d(x_0, y_0) \leq d(x_1, y_1) \leq \dots \leq d(x_{r-1}, y_{r-1}) \leq d(x_0, y_0).$$

Therefore every inequality is actually an equality and the reverse cycle  $\bar{c}$  is also in  $\mathcal{G}(W)$ . This proves that the graph  $\mathcal{G}(W)$  is cycle-reverter whenever  $W$  is metrizable. Conversely, if  $\mathcal{G}(W)$  is cycle-reverter then by Lemma 4.7 we can define in  $P = \mathcal{X}^{(2)}$  a colored poset structure where the order is induced by the graph  $\mathcal{G}_1(W)$  (i.e.  $v < w$  if there is a path from  $v$  to  $w$  in  $\mathcal{G}_1(W)$ ) and the connected components of  $\mathcal{G}_0(W)$  are monochromatic. By Proposition 4.2 we can construct a height function  $h$  for which every connected components of  $\mathcal{G}_0(W)$  is horizontal. In particular  $\Pr(x|y) < \Pr(x|z)$  if and only if  $h(\{x, y\}) < h(\{x, z\})$ . Hence, the function  $f(x, y) = h(\{x, y\})$  is symmetric and coherent with  $W$ , then  $W$  is metrizable (a metric can be constructed as in Equation (2)).  $\square$

**Remark 4.9.** When  $\mathcal{G}(W)$  is cycle-reverter we have the following algorithm to obtain a metric  $d$  matching to the channel  $W : \mathcal{X} \rightarrow \mathcal{X}$ .

- (1) Consider the colored poset in  $P = \mathcal{X}^{(2)}$  whose partial order is induced by  $\mathcal{G}_1(W)$  and the colouring is given by the connected components of  $\mathcal{G}_0(W)$ .
- (2) Construct a height function  $h$  for  $P$  as in the proof of Proposition 4.2 (see also Remark 4.3) for which every monochromatic set is horizontal.
- (3) Let  $N$  be the maximum value of  $h$ . A metric is given by

$$d(x, y) = \begin{cases} 3N - h(\{x, y\}) & \text{if } x \neq y \\ 0 & \text{if } x = y \end{cases}$$

**Example 4.10.** Let  $\mathcal{X} = \{0, 1, 2, 3\}$  and we consider the channel  $W : \mathcal{X} \rightarrow \mathcal{X}$  whose transition matrix is given by

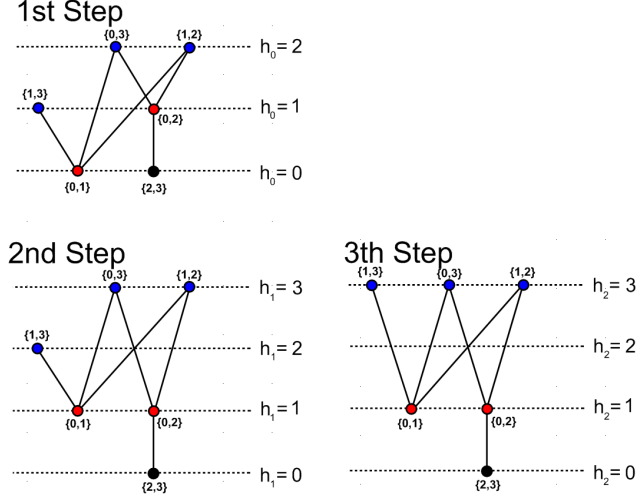
$$[W] = \begin{pmatrix} 2/3 & 0 & 0 & 1/3 \\ 0 & 1/2 & 1/4 & 1/4 \\ 1/6 & 1/3 & 1/2 & 0 \\ 1/4 & 1/4 & 0 & 1/2 \end{pmatrix}.$$

We can apply the steps given in the proof of Proposition 4.2 to obtain a Hasse diagram for the colored poset associated with  $W$  (defined in Lemma 4.7). This process is illustrated in Figure 3, after three steps we obtaining the height function  $h_2 : \mathcal{X}^{(2)} \rightarrow \mathbb{N}$  given by  $h_2(\{2, 3\}) = 0$ ,  $h_2(\{0, 1\}) = h_2(\{0, 2\}) = 1$  and  $h_2(\{1, 3\}) = h_2(\{0, 3\}) = h_2(\{1, 2\}) = 3$ . A metric matched to  $W$  is given by  $d(x, y) = 16 - h_2(\{x, y\})$  when  $x \neq y$  and 0 otherwise.

## 5. THE BAC CHANNEL IS METRIZABLE

We consider the  $n$ -fold BAC channel  $BAC^n(p, q)$  with parameters  $p, q \in [0, 1]$  and  $p + q < 1$ . The case  $pq = 0$  corresponds to the  $Z$ -channel which we know it is metrizable (Theorem 6 of [4]), then we can assume  $pq > 0$ . Each entry of its transition matrix  $M_n(p, q)$  is of the form

$$\Pr(x|y) = p^a(1-p)^b q^c(1-q)^d$$

FIGURE 3. Obtaining a metric matched to the channel  $W$ .

with  $a + b + c + d = n$  and  $a + d = w(x)$  (the Hamming weight of  $x \in \mathbb{Z}_2^n$ ). If we consider other word  $y' \in \mathbb{Z}_2^n$  with  $\Pr(x|y') = p^{a'}(1-p)^{b'}q^{c'}(1-q)^{d'}$ , since  $a + d = a' + d'$  and  $b + c = b' + c'$ , taking the quotient we have:

$$(4) \quad \frac{\Pr(x|y)}{\Pr(x|y')} = \left(\frac{1-p}{q}\right)^{d-d'} \cdot \left(\frac{1-q}{p}\right)^{d-d'}.$$

This identity will be useful in our proof of metrizability of the BAC channel.

**Lemma 5.1.** *Let  $W : \mathcal{X} \rightarrow \mathcal{X}$  be a channel. The graph  $\mathcal{G}(W)$  is cycle-reverter if and only if every sequence  $x_0, x_1, \dots, x_{r-1} \in \mathcal{X}$  ( $r \geq 3$ ) satisfying  $x_i \neq x_{i+1}$  and  $\Pr(x_i|x_{i-1}) \leq \Pr(x_i|x_{i+1})$  for  $i : 0 \leq i < r$  also satisfy  $\Pr(x_i|x_{i-1}) = \Pr(x_i|x_{i+1})$  for  $i : 0 \leq i < r$  (where the indices are considered modulo  $r$ ).*

*Proof.* We suppose that the graph  $\mathcal{G}(W)$  is cycle-reverter and consider a sequence  $x_0, x_1, \dots, x_{r-1} \in \mathcal{X}$  satisfying  $x_i \neq x_{i+1}$  and  $\Pr(x_i|x_{i-1}) \leq \Pr(x_i|x_{i+1})$  for  $i : 0 \leq i < r$ . Then we have a cycle  $c : v_0 \rightarrow \dots \rightarrow v_{r-1} \rightarrow v_0$  in  $\mathcal{G}(W)$  given by  $v_i = \{x_i, x_{i+1}\}$  for  $0 \leq i < r$ . Since this graph is cycle-reverter its reverse cycle  $\bar{c}$  is also a cycle in  $\mathcal{G}(W)$  which imply  $\Pr(x_i|x_{i-1}) = \Pr(x_i|x_{i+1})$  for  $i : 0 \leq i < r$ . Now we suppose that the graph  $\mathcal{G}(W)$  is not cycle-reverter and we will construct a sequence  $x_0, x_1, \dots, x_{r-1} \in \mathcal{X}$  satisfying  $x_i \neq x_{i+1}$  and  $\Pr(x_i|x_{i-1}) \leq \Pr(x_i|x_{i+1})$  for  $i : 0 \leq i < r$  where at least one inequality is strict. We consider a cycle  $c : v_0 \rightarrow v_1 \rightarrow \dots \rightarrow v_{r-1} \rightarrow v_r = v_0$  in  $\mathcal{G}(W)$  of minimal length  $r \geq 3$  whose reverse cycle  $\bar{c}$  is not in this graph. We remark that the fact that its reverse cycle  $\bar{c}$  is not in  $\mathcal{G}(W)$  is equivalent to the existence of some arrow in  $c$  which is also an arrow in  $\mathcal{G}_1(W)$ . The vertices in  $c$  are pairwise disjoint since otherwise we could take a sub-cycle of  $c$  containing some arrow of  $\mathcal{G}_1(W)$  contradicting the minimality of  $r$ . If for some  $i : 0 \leq i < r$  we have that  $v_i \cap v_{i+1} \cap v_{i+2} = \{x\}$ , then there exists  $y, z, t$  pairwise distinct such that  $v_i = \{x, y\}$ ,  $v_{i+1} = \{x, z\}$  and  $v_{i+2} = \{x, t\}$ . Thus  $\Pr(x|y) \leq \Pr(x|z) \leq \Pr(x|t)$  and the arrow  $v_i \rightarrow v_{i+2}$  is also in  $\mathcal{G}(W)$ . If

some of the arrows  $v_i \rightarrow v_{i+1}$  or  $v_{i+1} \rightarrow v_{i+2}$  is in  $\mathcal{G}_1(W)$  then  $v_i \rightarrow v_{i+2}$  is also in  $\mathcal{G}_1(W)$ , so we could substituting these two arrows for the last obtaining a new cycle, whose reverse is not in  $\mathcal{G}(W)$  and length  $r - 1$  which contradict the minimality of  $r$ . Therefore  $v_i \cap v_{i+1} \cap v_{i+2} = \emptyset$  for all  $i : 0 \leq i < r$  and there exists a sequence  $x_0, x_1, \dots, x_{r-1} \in \mathcal{X}$  such that  $v_i = \{x_i, x_{i+1}\}$ . This sequence satisfy  $\Pr(x_i|x_{i-1}) \leq \Pr(x_i|x_{i+1})$  for  $i : 0 \leq i < r$  with at least strict inequality (the corresponding to the edge in  $\mathcal{G}_1(W)$ ).  $\square$

**Theorem 5.2.** *Let  $n \geq 2$  and  $(p, q) \in [0, 1]^2$  with  $p + q < 1$ . Then, the channel  $W = BAC^n(p, q)$  is metrizable.*

*Proof.* By Theorem 4.8, it is enough to prove that its associated graph  $\mathcal{G}(W)$  is cycle-reverter. We assume, to the contrary, that this graph is not cycle-reverter and by Lemma 5.1 there exists a sequence  $x_0, x_1, \dots, x_{r-1} \in \mathcal{X}$  such that  $x_i \neq x_{i+1}$  and

$$(5) \quad \Pr(x_i|x_{i-1}) \leq \Pr(x_i|x_{i+1}), \quad \forall i : 0 \leq i < r,$$

where the indices are taken modulo  $r$  and where at least one of these inequality is strict. We write these conditional probability as  $\Pr(x_i|x_{i-1}) = p^{a_i}(1-p)^{b_i}q^{c_i}(1-q)^{d_i}$  for  $i : 0 \leq i < r$ . Therefore  $\Pr(x_{i-1}|x_i) = p^{c_i}(1-p)^{b_i}q^{a_i}(1-q)^{d_i}$  for  $i : 0 \leq i < r$  and applying Equation 4 we obtain

$$(6) \quad \begin{aligned} \frac{\Pr(x_i|x_{i+1})}{\Pr(x_i|x_{i-1})} &= \frac{p^{c_{i+1}}(1-p)^{b_{i+1}}q^{a_{i+1}}(1-q)^{d_{i+1}}}{p^{a_i}(1-p)^{b_i}q^{c_i}(1-q)^{d_i}} \\ &= \left(\frac{1-p}{q}\right)^{b_{i+1}-b_i} \left(\frac{1-q}{p}\right)^{d_{i+1}-d_i}, \end{aligned}$$

for  $0 \leq i < r$ . Multiplying these  $r$  inequalities we have

$$\begin{aligned} \prod_{i=0}^{r-1} \frac{\Pr(x_i|x_{i+1})}{\Pr(x_i|x_{i-1})} &= \prod_{i=0}^{r-1} \left(\frac{1-p}{q}\right)^{b_{i+1}-b_i} \left(\frac{1-q}{p}\right)^{d_{i+1}-d_i} \\ &= \left(\frac{1-p}{q}\right)^{\sum_{i=0}^{r-1} (b_{i+1}-b_i)} \left(\frac{1-q}{p}\right)^{\sum_{i=0}^{r-1} (d_{i+1}-d_i)} = 1 \end{aligned}$$

since  $\sum_{i=0}^{r-1} (b_{i+1}-b_i) = \sum_{i=0}^{r-1} (d_{i+1}-d_i) = 0$ . But by (5) this product is grater than 1 (since at least one inequality is strict) which is a contradiction. Therefore  $W$  is metrizable.  $\square$

## 6. CONCLUDING REMARKS AND FURTHER PROBLEMS

In this work we approach the problem of metrization for the  $n$ -fold BAC channels in the sense of definition used in [3] and [4]. A existence proof and an algorithm to construct a metric matching to the BAC channels are provided. An interesting problem would be to describe the set  $D_1(n, p, q)$  of metrics matching with the channel  $BAC^n(p, q)$ . This set is non-empty by Theorem 5.2 and closed by linear combination with positive real coefficients, so we could looking forward to a minimal generator for this set. Describing this set allows to choose good metrics according to a certain criterion. One possible criterion would be selecting according to how easy is to compute it. Other criteria could be according to how good the metric fit with the channel according to the following definition.

**Definition 6.1.** Let  $W : \mathcal{X} \rightarrow \mathcal{X}$  be a channel and  $d$  be a metric compatible to  $W$ . The metric  $d$  is matched to the channel  $W$  with order  $n$  if  $d^m$  is a metric matched to  $W^m$  for all  $m : 1 \leq m \leq n$ , where  $d^m : \mathcal{X}^m \rightarrow \mathcal{X}^m$  is given by

$$d^m(x, y) = \sum_{k=1}^m d(x_k, y_k)$$

and

$$\Pr_{W^m}(x|y) = \prod_{k=1}^m \Pr_W(x_k|y_k).$$

If  $d$  matched to  $W$  with order  $n$  for all  $n \geq 1$  we say that the metric  $d$  matches completely to  $W$ .

We also define the order of metrizable of  $W$  as the maximum  $n$  (possibly infinite) for which there exists a metric matched to  $W$  with order  $n$ . It would be interesting to determine the order of metrizable for the BAC channels or at least to determine which of these channels admit a matched metric with order  $n \geq 2$ . In [3] it was approached the problem of determining when a channel  $W : \mathcal{X} \rightarrow \mathcal{X}$  is completely metrizable for alphabets of lengths  $\#\mathcal{X} = 2, 3$  and proved that the channel  $BAC^1(p, q)$  has a matched metric with order  $\infty$  if and only if  $p = q$  (symmetric channel). In that case the Hamming metric matches completely to this channel as mentioned.

#### REFERENCES

- [1] J. C. Y. Chiang, J. K. Wolf, "On channels and codes for the Lee metric", *Inf. Control*, vol. 19, no. 2, pp. 159-173, 1971.
- [2] J. L. Massey, "Notes on coding theory", *Cambridge, MA, USA: MIT Press*, 1967.
- [3] G. S guin, "On metrics matched to the discrete memoryless channel", *J. Franklin Inst.*, vol. 309, no. 3, pp. 179-189, 1980.
- [4] M. Firer, J. L. Walker, "Matched metrics and channels", *IEEE Transactions on Information Theory*, vol. 62, no. 3, pp. 1150-1156, 2016.
- [5] R. G. L. D'Oliveira, M. Firer, "Channel metrization", *arXiv:1510.03104*, 2016.
- [6] A. Poplawsky, "On matched metric and channel problem", *ArXiv:1606.02763v1*, 2016.

INSTITUTE OF MATHEMATICS, STATE UNIVERSITY OF CAMPINAS, BRAZIL  
*E-mail address:* `cqureshi@ime.unicamp.br`